

Universidade Federal do Rio de Janeiro

**Instituto Tércio Pacitti de Aplicações e
Pesquisas Computacionais**

Marcio Marinho Baptista

**CODECS DE VÍDEO PARA TRANSMISSÃO
EM TEMPO REAL:**

**Descrição e Comparação dos
Padrões Atuais com os Novos**

Rio de Janeiro

2016

Marcio Marinho Baptista

**CODECS DE VÍDEO PARA TRANSMISSÃO EM
TEMPO REAL:**

**Descrição e Comparação dos Padrões Atuais
com os Novos**

Monografia apresentada para obtenção do título de Especialista em Gerência de Redes de Computadores no Curso de Pós-Graduação Lato Sensu em Gerência de Redes de Computadores e Tecnologia Internet do Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais da Universidade Federal do Rio de Janeiro – NCE/UFRJ.

Orientador:

Claudio Miceli de Farias, DSc, UFRJ, Brasil

Rio de Janeiro

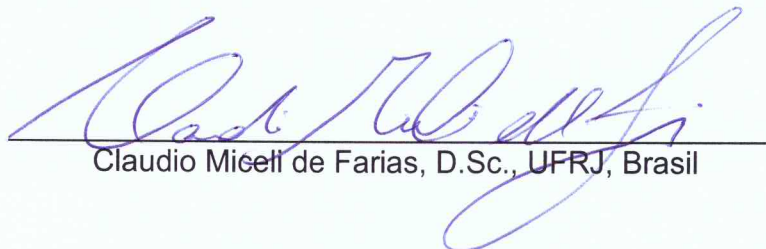
2016

Marcio Marinho Baptista

**CODECS DE VÍDEO PARA TRANSMISSÃO EM
TEMPO REAL:
Descrição e comparação dos padrões atuais
com os novos**

Monografia apresentada para obtenção do título de Especialista em Gerência de Redes de Computadores no Curso de Pós-Graduação Lato Sensu em Gerência de Redes de Computadores e Tecnologia Internet do Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais da Universidade Federal do Rio de Janeiro – NCE/UFRJ.

Aprovada em março de 2016.



Claudio Miceli de Farias, D.Sc., UFRJ, Brasil

Dedico o presente trabalho a minha Esposa e a minha Filha que com seu amor me deram a força necessária para a conclusão.

AGRADECIMENTOS

Agradeço a minha esposa Simone, que de forma especial e carinhosa me deu força e coragem, me apoiando nos momentos de dificuldades, quero agradecer também a minha filha Ana Carolina, que embora não tivesse conhecimento disto, mas iluminou de maneira especial os meus pensamentos me levando a buscar mais conhecimentos.

Agradeço ao professor Claudio Miceli. pela paciência na orientação e incentivo que tornaram possível a conclusão desta monografia.

Agradeço ao professor Moacyr, responsável pela coordenação do excelente curso de especialização.

O principal agradecimento que é a Deus, pela força e coragem durante toda esta longa caminhada.

RESUMO

BAPTISTA, Marcio Marinho. **CODECS DE VÍDEO PARA TRANSMISSÃO EM TEMPO REAL: Descrição e Comparação dos Padrões Atuais com os Novos.** Monografia (Especialização em Gerência de Redes e Tecnologia Internet). Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2016.

O trabalho a seguir aborda os *codecs* para transmissão de vídeo em tempo real. Ao decorrer do mesmo, serão enunciadas e explicadas as principais dificuldades na rede de computadores para a transmissão desses vídeos. As principais descrições e evoluções das normas de codificação e compressão como, por exemplo: a norma *HEVC* conhecida também como *codec* H.265. Foi feita uma descrição do novo padrão *HEVC* (H.265), contendo as melhorias apresentadas nesse novo padrão. Ocorrerá uma comparação do novo padrão H.265, que obteve uma redução da utilização de banda na transmissão e recepção de vídeos de aproximadamente 40% em relação ao seu antecessor, o *codec* H.264, e será abordado também as evoluções do *codec* H.265, o que tem evoluído em relação ao desenvolvimento de *codecs* para transmissão de vídeos em 3D e a possibilidade de um aumento da qualidade e de uma maior sensação de realidade na transmissão e recepção dos vídeos em tempo real.

ABSTRACT

BAPTISTA, Marcio Marinho. **VIDEO CODECS TRANSMISSION IN REAL TIME:** Description and Comparison of Current Standards with the New. Monograph (specialization in Network Management and Internet Technology) Institute Tércio Pacitti Applications and Research Computing, Federal University of Rio de Janeiro Rio de Janeiro, in 2016.

The following work deals with the codecs for real-time video transmission. In the course of it, will be set out and explained the main difficulties in the computer network for the transmission of these videos. The main developments and descriptions of coding and compression standards such as, for example, known also as HEVC standard H.265 codec. A description of the new standard HEVC (H.265) containing the improvements presented in this new standard was made. There will be a comparison of the new standard H.265, which obtained a reduction of bandwidth utilization in the transmission and reception of video approximately 40% compared to its predecessor, the H.264 codec, and will also be addressed developments H. 265 codec, which has evolved for the development of codecs for transmission of 3D video and the possibility of increased quality and a greater sense of reality in the transmission and reception of video in real time.

LISTA DE FIGURAS

	Página
Figura 2.1 - Comparação entre latência e <i>jitter</i>	19
Figura 2.2 - Definição do <i>Skew</i> entre médias diferentes	20
Figura 2.3 - Transmissão Multimídia	21
Figura 2.4 - Estrutura de um Datagrama	23
Figura 2.5 - Modelo OSI	24
Figura 3.5 - Construção do <i>bitstream</i> por variante em zig-zag	35
Figura 4.6 - Arquitetura de blocos do decodificador <i>HEVC</i>	38
Figura 4.7 - Divisão recursiva da UC	39
Figura 5.8 - Ilustração da divisão de uma CTU em CUs	43
Figura 5.9 - PSNR x <i>BitRate</i> – <i>BlowingBubbles</i> – 416x240 PSNR x <i>BitRate</i> – <i>BasketballDrill</i> – 832x480	44

LISTA DE TABELAS

Tabela 5.2 – Comparação de desempenho entre *Codec* H.264 e H.265

Página
45

LISTA DE ABREVIATURAS E SIGLAS

CB	Bloco de Codificação
Codec	Codificação e Decodificação
CTB	Bloco de Árvore de Codificação
CTU	Unidade de Árvore de Codificação
CU	Unidade de Codificação
DTC	Transformada Discreta de Cos-seno
HD	Alta Definição
HE	Alta Eficiência
HECV	Alta Eficiência de Codificação de Vídeo
I	Modo Intra
IP	Protocolo de Internet
JITTER	Variação Estatística do Retardo
Kbps	Kilobytes por Segundo
LC	Baixa Complexidade
LD	Baixo Atraso
Mbps	Megabytes por Segundo
OSI	Organização Internacional para a Normalização
P2P	Pessoa para Pessoa
PB	Bloco de Predição
QoE	Qualidade de Experiência
QoS	Qualidade do Serviço
RA	Acesso Aleatório
RAM	Memória de Acesso aleatório
RPT	Protocolo de Transporte em Tempo Real
RTCP	Protocolo de Controle em Tempo Real
TB	Bloco da Transformada
UC	Unidade de Codificação
UHD	Ultra-Alta Definição
UP	Unidade de Predição
UT	Unidade de Transformada
VLC	Método de Redução na Taxa de Bits

SUMÁRIO

	Página
1 INTRODUÇÃO	12
2. TRANSMISSÃO MULTIMÍDIA NAS REDES DE COMPUTADORES	15
2.1 TRANSMISSÃO DE VÍDEO NAS REDES DE COMPUTADORES	15
2.2 TRÁFEGO DE VÍDEO	16
2.2.1 Latência	17
2.2.2 Jitter	18
2.2.3 Skew	19
2.3 QUALIDADE DA EXPERIÊNCIA QoE	20
2.4 QUALIDADE DO SERVIÇO QoS	21
2.5 TRANSMISSÃO MULTIMÍDIA	21
2.6 PROTOCOLOS DE TRANSFERÊNCIA	22
2.6.1 RTP	22
2.6.2 RTCP	23
2.7 STREAMING DE VÍDEO	25
2.7.1 Buffer	26
3. OS CODECS E A SUA RELEVÂNCIA	28
3.1 CONCEITOS BÁSICOS DE <i>CODECS</i>	28
3.2 CODIFICAÇÃO DE VÍDEO	29
3.2.1 Redundâncias Espaciais ou Temporais	30
3.2.2 Redundância Estatística	30
3.2.3 Exploração do Sistema Visual Humano	30
3.3 <i>CODEC</i> H.264	31
3.4 <i>CODEC</i> H.265	32
3.5 EFICIÊNCIA DOS <i>CODECS</i>	33
3.5.1 Compensação de Movimento	33
3.5.2 Transformada Discreta de Co-seno (DCT)	34
3.5.3 Quantificação e Codificação por VLC	35
3.5.4 Organização do <i>Bitstream</i>	36
4. O PADRÃO HEVC	37
4.1 Conceito de <i>HEVC</i> e suas Características	37
4.2 Estrutura do Codec <i>HEVC</i>	38
4.2.1 Unidade de Codificação (UC)	39
4.2.2 Unidade de Predição (UP)	40
4.2.3 Unidade de Transformada (UT)	40
4.2.4 Unidade Árvore de Codificação (CTU)	40
4.3 Quadros	41
5. COMPARAÇÃO DO H.265 COM O SEU ANTECESSOR	42
5.1 Comparação dos <i>Codecs</i> de Padrões Atuais com os Novos	42
5.2 Condições de Restrição	43
5.2.1 <i>Random Access</i> (RA)	43
5.2.2 <i>Low Delay</i> (LD)	44
6. PERSPECTIVAS DE EVOLUÇÃO DO H.265	47
6.1 O Potencial do H.265	47
7. CONCLUSÃO	50
8 TRABALHOS FUTUROS	52
REFERÊNCIAS	53

1. INTRODUÇÃO

Em 2014, aplicações que transmitem vídeo em tempo real na internet atingiram 64% do tráfego de rede e com previsão de chegar até 80% no final de 2019. Este percentual é na maioria para streaming de vídeo armazenado e streaming de vídeo em tempo real, Por exemplo: Youtube, NetFlix entre outros, não incluindo trocas de vídeos através de P2P (peer-two-peer) (CISCO, 2014).

O termo codec provém de uma abreviatura do conjunto de palavras (Codificador e Decodificador) e como seu nome indica, consiste numa especificação de software que permite comprimir e descomprimir os arquivos. Dessa forma, a demanda de codecs mais eficientes se faz necessário devido ao crescente aumento e fácil acesso de links de internet de arquivos de mídias. A transmissão de vídeo em tempo real consome uma largura de banda significativa.

As limitações da infraestrutura de rede e a inexistência de codecs capazes de compactar e descompactar os arquivos de mídia transmitidos, mantendo a qualidade do seu conteúdo original. Estes são os fatores que exigem uma análise mais equilibrada para conceituar, e idealizar soluções inovadoras para o desenvolvimento de codecs mais eficientes capazes de permitir altas taxas de compactação com mínima perda de qualidade na transmissão e descompactação do arquivo de mídia.

Os conceitos de transmissão de vídeo nas redes de computadores envolvem também outros processos internos relevantes tais como: Latência, Jitter, Skew. Todos eles dentro de um contexto organizado obedecendo aos parâmetros dos Protocolos de Transferência; RTP e RTCP, úteis para transporte de dados em tempo real, o uso de *Streaming* de vídeo e do *Buffer* no processo de transmissão de

multimídia. Este processo complexo de transmissão de mídia em tempo real segue a uma sequência lógica organizada, no qual cada um dos elementos acima citados, assim como também a infraestrutura de rede, os protocolos de transmissão e os recursos de compactação e descompactação da informação participantes desse processo tem seu papel importante,

Foi descrita a estrutura dos *codecs* a partir dos princípios da análise temporal e/ou espacial das imagens primárias, considerando as redundâncias das imagens encontradas quadro a quadro, abordando a eficiência dos *codecs* e as deficiências da visão humana, em detalhe descrevemos as diferentes técnicas na eliminação de redundância na codificação de vídeo.

Descrevemos a arquitetura do padrão *HEVC* destacando as novas funcionalidades, e as operações típicas para a produção do fluxo de bits válido. Sendo assim, citamos a estrutura hierárquica das unidades e blocos que compõem o *codec HEVC*. Portanto, neste contexto, além da exploração da transmissão de vídeo em tempo real será realizado um estudo de fatores importantes no dimensionamento do melhor cenário, garantindo uma qualidade boa para os vídeos transmitidos.

O trabalho tem como objetivo analisar e comparar o desempenho do *codec* H.265, assim como do seu antecessor o *codec* H.264, apresentando os elementos que compõem tais *codecs* e seus mecanismos de compactação, analisamos também diversos estudos e publicações relacionadas à transmissão de vídeo em tempo real e os processos envolvidos no mesmo.

A pesquisa foi feita de forma organizada, estudando desde os elementos mais simples, até os mais complexos envolvidos na transmissão e compactação de vídeo, definindo bem cada um dos processos e elementos relacionados ao assunto.

Dessa forma, o aumento da utilização de vídeos na internet, o surgimento de novas aplicações, assim como videoconferência em tempo real, requerem uma infraestrutura de rede mais adequada à crescente demanda, os problemas na rede que geram uma transmissão de vídeo com falhas são detalhados no capítulo 2.

No capítulo 3 serão abordados os *codecs* de vídeo e sua importância, atualmente existe uma crescente demanda de serviços de *streaming* de vídeos, sobrecarregando o tráfego da internet, para isso a necessidade da eficiência de um *codec* que atenda essa demanda com taxas de compressão elevadas ocupando menor largura de banda da internet.

No capítulo 4 será explicado como funciona o novo padrão *HEVC* (*High Efficiency Video Coding*), também conhecido como H.265, esse padrão de *codec* foi a princípio desenvolvido com base em duas configurações de codificação: Alta Eficiência (HE) e Baixa Complexidade (LC) (BELTRÃO, 2011).

Para melhor entendimento das vantagens do novo padrão *HEVC*, foi feita uma comparação do novo padrão com o seu antecessor, o padrão H.264.

São descritas as características de cada padrão no capítulo 5, e porque o *codec* H.265 chega a ter uma economia de banda de até 40% em relação ao conteúdo codificado pelo H.264 (OLIVEIRA, 2014).

Com relação à evolução do *codec* H.265, no capítulo 6, são descritos os estudos que estão sendo feitos para adicionar extensões ao *HEVC* para transmissão de vídeos em 3D (GOPALAKRISHNA, 2013). As conclusões do trabalho e as referências bibliográficas são apresentadas em seguida.

2. TRANSMISSÃO MULTIMÍDIA NAS REDES DE COMPUTADORES

A finalidade deste capítulo é abordar alguns conceitos de transmissão de vídeo nas redes de computadores, assim como os elementos envolvidos: *Latência*, *Jitter*, *Skew*. Citamos também os Protocolos de Transferência: RTP e RTCP, utilizados para transportar dados em tempo real, o uso de *Streaming* de vídeo, da mesma forma o papel relevante do *Buffer* no processo de transmissão de vídeo pela rede de computadores.

2.1 TRANSMISSÃO DE VÍDEO NAS REDES DE COMPUTADORES

Quando se transmite vídeos pela rede de computadores e eles não estão sincronizados com o áudio, o processo de exibição para os ouvidos e olhos humanos podem facilmente perceber tal falha; além da latência, congestionamentos na rede também têm efeitos sérios sobre o tráfego de tempo real (TSCHÖKE, 2001).

Se a rede estiver congestionada, o efeito sobre um tráfego que não seja em tempo real, será a transferência que levará mais tempo para completar. Entretanto, vídeos em tempo real se tornam obsoletos e serão descartados se não chegarem sincronizados no momento exato da transmissão. Se nenhuma ação apropriada for tomada, a retransmissão dos pacotes perdidos agravaria a situação.

Apenas aumentar a largura de banda não resolveria o problema de transmissão em rajadas. Para a maioria das aplicações multimídia, o receptor tem um *buffer* de tamanho limitado. Se nenhuma medida for tomada para regular o fluxo de dados, ele pode gerar uma sobrecarga (ou um fluxo leve demais) no *buffer* da aplicação. Quando os dados chegarem muito rápido, o *buffer* irá sobrecarregar-se e alguns pacotes serão perdidos, resultando em uma qualidade muito ruim (TSCHÖKE, 2001).

Uma das principais aplicações corporativas para comunicação de vídeo em tempo real, a videoconferência, por exemplo, é sensível a atrasos e perdas de pacotes (ALMEIDA, 2004). Nesse caso, em geral a infraestrutura da rede IP deve suportar um *jitter* $\leq 30\text{ms}$ e uma perda de pacotes $\leq 1\%$. Já para tele presença, a rede IP deve tratar um tráfego ainda mais sensível, com *jitter* $\leq 10\text{ms}$, a perda de pacotes $\leq 0,05\%$ e a grande necessidade de banda (HERSENT, 2002).

Já as aplicações de vídeo IP para segurança patrimonial, tratam-se basicamente de *streaming* com algumas necessidades especiais e o total de banda necessária depende da quantidade de visualizações simultâneas.

A infraestrutura de rede, por sua vez, deve ser pensada de forma que a latência seja $\leq 150\text{ms}$, o *jitter* $\leq 10\text{ms}$, a perda de pacotes $\leq 0,05\%$ e os requerimentos de banda podem variar de 200kbps a até 3,5 Mbps (CISCO, 2014).

2.2 TRÁFEGO DE VÍDEO

Em relação apenas a tráfego de vídeo *streaming* e sob-demanda, os requerimentos de rede são mais leves, com uma boa infraestrutura de rede o *delay* e *jitter*, na maior parte das vezes, não são problemas. Porém, a perda de pacotes torna-se um requerimento importante e a rede IP deve suportar perdas $\leq 0,05\%$ para tráfego HD (*high definition*), por exemplo (CISCO, 2014).

Considerando os padrões de tráfego mencionados anteriormente conclui-se que numa arquitetura de rede IP ideal, deve-se controlar: o retardo fim-a-fim, a taxa de perda de pacotes, a variação do atraso fim-a-fim (*jitter*) e a vazão (DANTAS, 2006).

Na transmissão de multimídia na rede podemos observar três conceitos envolvidos:

- Latência
- *Jitter*
- *Skew*

2.2.1 Latência

Para Brun (2002), a Latência é o tempo que um pacote leva desde origem até o seu destino final. A conversação através da rede se prejudica, caso esse atraso seja muito grande, o que dificulta a interatividade necessária na transmissão do pacote.

Na opinião de Roesler (2001), os principais responsáveis pela latência são o atraso de transmissão, de codificação e de empacotamento, que podem ser explicados como:

- Atraso de transmissão: tempo que leva para o pacote sair da placa de rede do computador origem e chegar na placa de rede do computador destino. Esse tempo envolve uma série de fatores, como por exemplo:
- Atraso no meio físico: é o atraso de propagação da mensagem no meio de transmissão, e varia bastante.
- Atraso de processamento nos equipamentos intermediários, como roteadores e *switches*.
- Atraso devido ao tempo de espera nas filas de transmissão dos equipamentos intermediários: esse valor depende do congestionamento da rede no momento, e varia bastante, dependendo do tamanho da fila. Quanto menor a fila, menor o atraso, mas aumenta a probabilidade de descarte do pacote no caso de congestionamento.

- Atraso de codificação e decodificação: tempo de processamento na máquina origem na máquina destino para codificação e decodificação de sinais, respectivamente.
- Atraso de empacotamento e desempacotamento: depois de codificado, o dado deve ser empacotado através dos níveis na pilha de protocolos a fim de ser transmitido na rede.

Além disso, dependendo do *jitter* da transmissão, a aplicação de tempo real deverá criar um buffer para homogeneizar a entrega de pacotes ao usuário, criando um novo atraso no sistema (ROESLER, 2001).

2.2.2 Jitter

Segundo Vasconcelos (2010), o *jitter* é a variação do retardo. Prejudicará as transmissões do tipo streaming, nas quais os dados são exibidos imediatamente após a recepção, com pouca ou nenhuma bufferização. A variação do retardo fará com que o vídeo recebido apresente pequenas pausas. Uma solução é usar um buffer maior para compatibilizar a chegada dos dados, com taxa variável, com a exibição a uma taxa constante. O problema aqui é que o uso do buffer maior também resultará em aumento do retardo, apesar de permanecer constante.

O *jitter* ou variação do atraso fim-a-fim mostra as variações no tempo de entrega de um pacote em um determinado tempo. Essa variação tem relação direta com a demanda de uso da rede (DA SILVA, 2005).

O mesmo contabiliza o atraso total sofrido por um pacote até atingir seu destino final a rede. Este parâmetro de medida leva em conta o tempo que os

pacotes esperam em filas de roteamento, o tempo de propagação no meio físico e o tempo gasto com processamento nos terminais de rede.

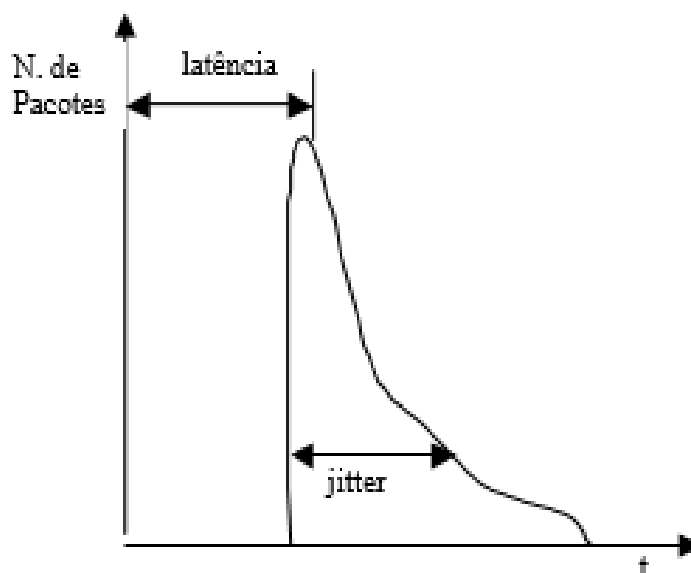


Fig. 2.1: Comparação entre latência e *jitter*

Fonte: (Roesler, 2001)

2.2.3 Skew

O *skew* é um parâmetro utilizado para medir a diferença entre os tempos de chegada de diferentes mídias que deveriam estar sincronizadas. Desta forma, comparamos o tempo de sincronia do áudio e vídeo numa aplicação.

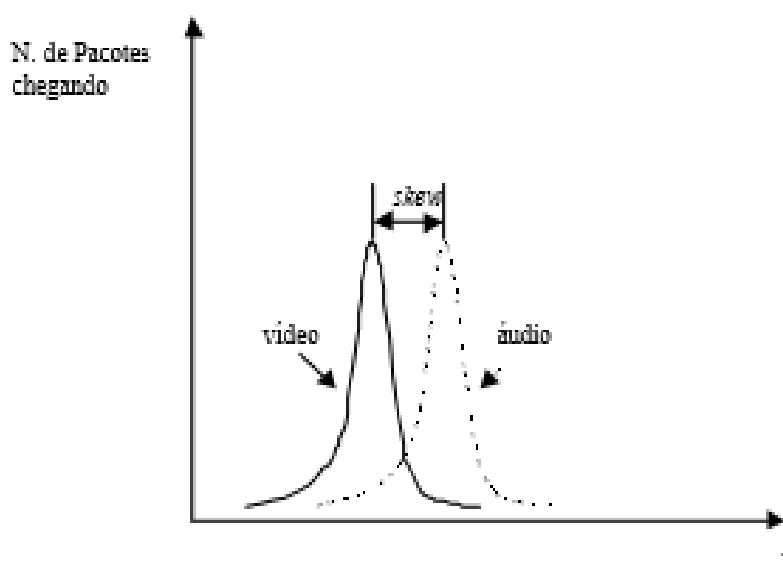


Fig. 2.2: Definição do Skew entre médias diferentes

Fonte: (Roesler, 2001)

A taxa de perda de pacotes é a razão entre o número de pacotes recebidos e o número de pacotes enviados em certo intervalo de tempo. Este parâmetro pode indicar que a taxa de transmissão de um terminal está incompatível com a capacidade da rede (SILVA, 2005).

2.3 QUALIDADE DA EXPERIÊNCIA (QoE)

Atendidos os requisitos acima, é possível dar aos usuários a qualidade de experiência (QoE) que aplicações de vídeo demandam. O QoE (*Quality of Experience*) procura mensurar algo muito mais subjetivo como a satisfação do usuário, sintetiza-se em sentimentos humanos como 'bom', 'excelente', 'pobre', etc.

2.4 QUALIDADE DO SERVIÇO (QoS)

Por outro lado, a QoS (*Quality of Services*) avalia a qualidade do serviço, baseado no tráfego da sua rede definindo prioridades e limites de forma a melhorar a percepção do usuário (Teleco, 2007).

2.5 TRANSMISSÃO MULTIMÍDIA

Na transmissão multimídia na rede de computadores se encontram subdivididas em duas partes:

- Teleconferência (carrega interatividade)
- Transmissão unidirecional (Transmissão de um único emissor e vários receptores).

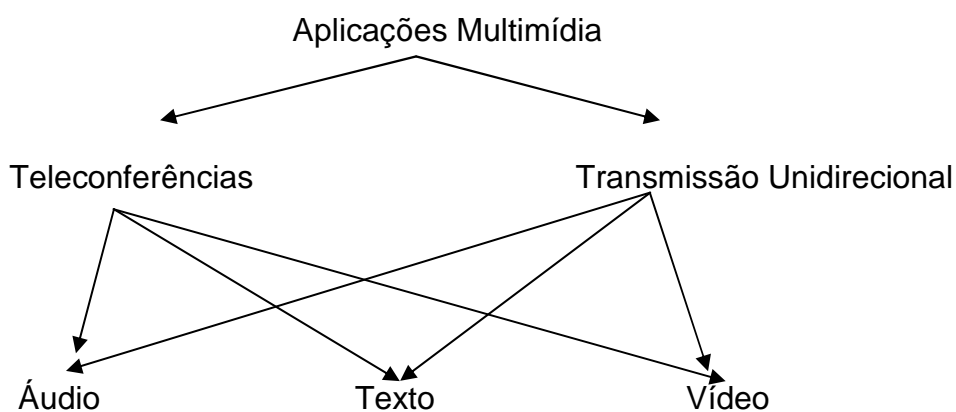


Fig. 2.3: Transmissão Multimídia

Fonte: (Roesler, 2001)

2.6 PROTOCOLOS DE TRANSFERÊNCIA

Por outro lado, Para transportar dados multimídia em tempo real, são necessários protocolos capazes de levar consigo sincronia e tempo, neste caso nos referimos ao RTP (*Real-time Transport Protocol*) e RTCP (ROESLER V 2009).

2.6.1 RTP

O RTP é utilizado para transportar dados em tempo real, e utiliza o RTCP para monitorar a qualidade de serviço e levar informações sobre os participantes de uma sessão em andamento, como por exemplo, uma conferência de áudio entre diversos participantes.

Algumas das funções do RTP:

O RTP fornece um feedback sobre a qualidade de serviço obtida na distribuição de dados RTP. Exemplos de utilização são: controle de codificadores adaptativos (muda algoritmo de compactação dependendo da qualidade), diagnóstico de problemas na rede, e outros;

Um transmissor com múltiplas sessões RTP (áudio e vídeo) utiliza um SSRC para cada sessão, e o receptor precisa um nome canônico para identificar a origem e poder sincronizar as sessões;

Encarrega-se também de controlar a periodicidade de envio dos pacotes RTCP;

Um servidor de mídia contínua encapsula um pedaço da mídia num pacote RTP, depois esse pacote é encapsulado num segmento UDP e então entregue ao IP. O lado cliente extrai do UDP o pacote RTP, de onde extrai o pedaço da mídia e passa para o media player. O RTP não provê mecanismos para certificar de que o dado chegará dentro da data prevista ou outra qualidade de serviço. Nem garante a

ordem da chegada dos dados, e nem mesmo a entrega deles. Os roteadores não têm como saber se um datagrama IP contém ou não um encapsulamento RTP. (KUROSE E ROSS, 2001)

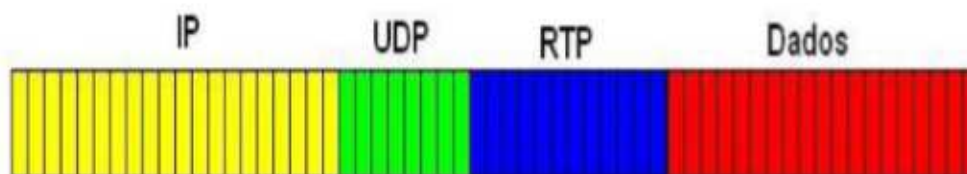


Fig. 2.4 Estrutura de um Datagrama

Fonte (KUROSE E ROSS, 2001).

Certamente o fato das aplicações RTP serem feitas na maioria das vezes em UDP/IP não significa que RTP requer UDP e IP. Ao passar para camada de transporte a mídia não comprimida é alocada em buffers que produzirão frames comprimidos de acordo com o algoritmo de compressão definido, esses frames são carregados nos pacotes RTP para envio e caso sejam muito grande eles poderão ser fragmentados em vários pacotes. Após o pacote ter sido enviado o transmissor não poderá descartar esses dados, pois talvez seja necessário para possíveis correções de erros.

2.6.2 RTCP

O protocolo RTCP (*Real-time Control Protocol*) tem o propósito de fornecer *feedback* sobre a qualidade de serviço obtida na distribuição de dados RTP, e consegue isso através de transmissões periódicas de pacotes de controle a todos participantes da sessão RTP, utilizando o mesmo mecanismo de distribuição do RTP (*unicast* ou *multicast*), possuindo uma porta específica de controle na sessão.

O intervalo mínimo sugerido entre pacotes RTCP é de 5 segundos (para evitar excesso de pacotes RTCP), mas numa sessão, o intervalo pode ir de 2 a 5 minutos. Um receptor deve desconsiderar um participante da estatística caso ele não se manifeste em 30 minutos

Um pacote RTCP é transmitido após o tempo calculado vezes um tempo randômico entre 0,5s e 1,5s. Isso é para evitar sincronização de pacotes entre várias entidades (transmissores e receptores), o que ocasionaria uma rajada de tráfego RTCP naquele momento.

Ambos protocolos se encontram contidos acima do nível 4 do modelo OSI

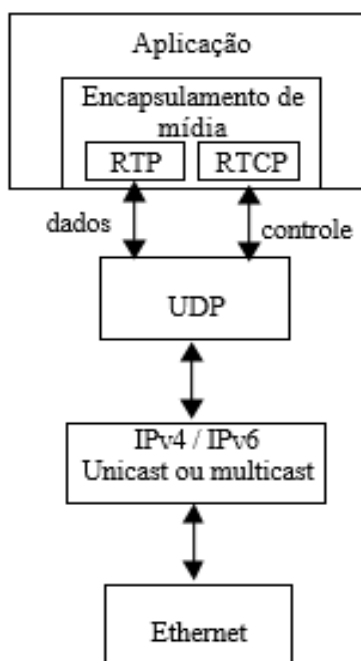


Fig. 2.5: Modelo OSI

Fonte: (Roesler, 2001)

Cinco tipos de pacotes RTCP são definidos pelo RTP:

- RTCP RR (Report Receiver RTCP – Relatório do Receptor RTCP) – O relatório estatístico que é um pacote RTCP é gerada a cada corrente RTP recebida e então o receptor envia àquele pacote para a árvore multicast a qual ele esta conectado.
- RTCP SR (Sender Report RTCP – Relatório do Transmissor RTCP) – São também informações estatísticas, contudo são enviadas pelo transmissor.
- RTCP SDES (Description of the Source RTCP – Descrição da Fonte RTCP) – são pacotes que contém identificação do participante, localização, e-mail etc.
- RTCP BYE (Gerenciamento de Membros) – esse pacote é gerado quando um participante deixa a sessão.
- RTCP APP (Applications Packages RTCP – Aplicações de Pacotes RTCP) – permite extensões definidas pela aplicação.

2.7 STREAMING DE VÍDEO

Novos desenvolvimentos tecnológicos permitiram a disponibilização de vídeos transmitidos pela internet na forma de *Streaming* de vídeo, esta tecnologia permite transmissão em tempo real do vídeo, possibilitando a inicialização da apresentação enquanto ocorre sua transmissão. Neste caso, o receptor solicita a transmissão de um determinado vídeo, o servidor inicia a transmissão dividindo o arquivo em pacotes, tais pacotes serão remontados em trechos (*Streams*) pedaços que podem ser visualizados individualmente sem necessidade de ter visualizado trechos

anteriores. Assim, esta tecnologia no presente momento está consolidada como uma solução à transmissão de vídeo em tempo real na internet.

Podemos indicar o funcionamento deste processo da seguinte forma: Tanto as imagens do vídeo assim como os áudios capturados por dispositivos digitais são comprimidos e colocados nos servidores, por solicitação do cliente. O servidor de *streaming* leva a transmissão para uma aplicação que tem como função adaptar a taxa de transmissão a realidade do canal de transmissão, a informação é empacotada pelos protocolos para seu transporte pela rede IP, a máquina receptora finalmente faz a aplicação de leitura e recebe os *streams*, sincronizando o áudio e o vídeo, ordenando e levando para o decodificador (*decoder*) para visualização final.

Alguns elementos envolvidos neste processo são fundamentais para garantir uma boa transmissão *streaming* de vídeo; Largura de banda, Atrasos, Perdas de pacotes, Decodificação eficiente (SILVA, 2005).

Cada um destes elementos tornam-se necessários e cada vez mais estão em desenvolvimento constante para melhorar a eficiência da transmissão em tempo real. Precisamos garantir um mínimo de largura de banda, como isso não acontece, necessitamos de um controle de congestionamento nos períodos críticos no tráfego de rede, o mesmo é feito pelo próprio *streaming* limitando a taxa de transmissão, o atraso inerente ao processo de transmissão requer no receptor final de um *Buffer* (retentor).

2.7.1 Buffer

Os *Buffers* são implementados em hardware e principalmente em software, que normalmente usam a memória RAM mais rápida para armazenar dados

temporários, devido ao tempo de acesso muito mais rápido comparado com as unidades de disco rígido.

As perdas de pacotes são inerentes à rede o que implica que o protocolo de *streaming* tem de responder de forma eficiente a tais perdas.

Assim, o papel dos decodificadores tornam-se relevantes, para manter a qualidade de áudio e vídeo, tal descodificação terá de ser simples, usando o mínimo de recursos do receptor final.

3. OS CODECS E A SUA RELEVÂNCIA

Neste capítulo será abordada a relevância dos *codecs* no processo de transmissão de vídeos, assim como a estrutura de funcionamento a partir dos princípios da análise temporal e/ou espacial das imagens primitivas e as redundâncias das imagens que a compõem quadro a quadro, assim como também, as vantagens da compressão de vídeo. Foi Analisada a eficiência dos *codecs* aproveitando as deficiências da visão humana, descrevendo detalhadamente as diferentes técnicas na eliminação de redundância na codificação de vídeo.

3.1 CONCEITOS BÁSICOS DE CODECS

Codecs são programas que codificam e decodificam arquivos de mídia, favorecendo compactação para armazenamento e descompactação para visualização. Em geral, diferentes *codecs* existentes hoje são utilizados para comprimir arquivos de vídeo, áudio e imagens, a fim de garantir que eles possuam um tamanho inferior à sua fonte original, para permitir que dessa forma, eles sejam mais fáceis de manusear, trabalhar e transmitir tais arquivos de mídia.

A maioria dos arquivos de dados multimídia contém áudio, vídeo e referências para a sincronização de ambos, de modo que a aplicação dos *codecs* apropriados devem ser obrigados a reproduzir um filme perfeitamente sincronizado no receptor final (ALMEIDA, 2004).

Geralmente são utilizados *codecs* que comprimem com uma certa perda de informação, porque a função fundamental da utilização de *codecs* é reduzir significativamente o tamanho do arquivo para uso mais fácil.

Normalmente, quando um arquivo de vídeo é convertido, tem lugar um processo automático, no qual o *codec* faz a parte do corte das informações sobre a imagem ou som original, e em seguida redefine um espaço menor, e finalmente o codifica.

O *codec* tem uma importância muito grande na transmissão de áudio e vídeo, permitindo que o armazenamento e transmissão sejam feitos com menores custos em termos de largura de banda e consumo de energia.

As aplicações em tempo real são críticas porque o *codec* sofre limitações no seu desempenho, assim como a velocidade da internet do usuário final. Isso pode ser refletido na taxa de compressão, ou seja, pode ser possível obter uma taxa de compressão não tão pequena quanto seria desejado. Teríamos que escolher entre aceitar o aumento do tráfego na rede ou reduzir a qualidade do vídeo (BELTRÃO, 2011).

3.2 CODIFICAÇÃO DE VÍDEO

Para Almeida, (2004) a codificação de vídeo é feita a partir da análise temporal e/ou espacial das imagens primitivas. A sua função é tirar proveito de algumas características do sinal de vídeo que permitem subtrair informação redundante, pouco importante ou mesmo tirar partido do funcionamento da visão humana. Tais propriedades seriam:

- Redundâncias espaciais ou temporais
- Redundância estatística
- Exploração do sistema visual humano

3.2.1 Redundâncias Espaciais ou Temporais

Dentro de uma imagem a quantidade de *pixels* muitas das vezes é quase do mesmo valor em relação a imagem anterior e posterior com isso teríamos que criar sistemas de predição entre imagens consecutivas de tal forma que seja apenas necessário enviar para cada uma, a diferença em relação à anterior, dessa forma nos permitiria reduzir a quantidade de informação a ser transportada na rede.

3.2.2 Redundância Estatística

Da mesma forma como acontece na codificação de Huffman, podem ser usados códigos mais curtos em relação aos que comumente vem aparecendo.

A codificação de Huffman é um método de compactação que usa as probabilidades de ocorrência dos símbolos no conjunto de dados a ser compactado para determinar códigos de tamanho variável para cada símbolo (COSTA, 2009).

3.2.3 Exploração do Sistema Visual Humano

As limitações do ângulo de visão do ser humano, e a limitada resolução espacial, permitem que o tamanho das imagens sejam apresentadas de diferentes formas, tais que o olho humano não percebe certas distorções.

Assim, como o número de imagens por segundo a serem captadas pelo olho humano não tem a capacidade de detectar variações rápidas de imagens, também em relação à qualidade de imagem que varia a percepção de pessoa a pessoa.

A visão humana é mais sensível à luminância do que à crominância, várias normas seguem uma estrutura em que a componente da primeira é quatro vezes superior à da segunda, pois as amostras de luminância são duas vezes mais na horizontal e na vertical da imagem.

3.3 CODEC H.264

É um *codec* flexível, que é usado por serviços, videoconferências, aplicações móveis, fornecedores de satélite de streaming, e para discos *Blue-ray*. Ele está muito bem escalado desde que foi apresentado pela primeira vez e é capaz de lidar com 3D, 48-60 fps e até 4K.

O padrão de disco *Blue-ray* não inclui atualmente provisões para algumas destas tecnologias, mas o próprio codec H.264 é capaz de lidar com eles.

Logo, o H.264/AVC, desenvolvido entre 1999 e 2003, que, posteriormente, fora estendido de várias outras formas de 2003 à 2009. O padrão H.264/AVC é uma tecnologia que possibilita aplicações de vídeo digital em áreas que não eram antes cobertas pelos seus antecessores, entre eles, o MPEG-2. Ele tem sido usado em transmissões de televisão em alta definição por satélite, cabo e difusão terrestre, sistemas de edição e aquisição de vídeo, filmadoras, aplicações de segurança, Internet, vídeo para dispositivos móveis, discos Blue-ray e aplicações de teleconferência ao vivo. No entanto, uma crescente diversidade de serviços, a popularidade crescente do vídeo em altas definições e o surgimento de resoluções além da alta definição (Ultra High Definition – UHD) estão criando uma demanda de eficiência de codificação superior a eficiência alcançada pelo H.264/AV.

Entretanto, o problema com o H.264, é que, enquanto ele pode lidar com esses tipos de codificações, não pode fazê-lo mantendo simultaneamente baixos os tamanhos dos arquivos (OLIVEIRA, 2014).

Algumas evoluções do H.264

- Compensação de movimento com tamanhos variáveis dos macro blocos: Desde o típico 16x16 até o mínimo de 4x4 com uma série de possibilidades distintas entre eles;

- A DCT é aplicada a blocos 4x4 e não a 8x8 como acontecia antes para uma codificação mais precisa; no entanto, são conferidos graus de liberdade para que possa ir até 16x16 na crominância ou 8x8 na luminância se tal for vantajoso;
- Menor complexidade na implementação do processo de precisão dos movimentos até $\frac{1}{4}$ de pixel;
- Utiliza coeficientes inteiros na DCT (ALMEIDA, 2004).

3.4 CODEC H.265

O *codec* H.265 foi projetado para utilizar substancialmente menos largura de banda graças às técnicas de codificação avançadas e um modelo de codificação / descodificação mais sofisticado.

Ao contrário de H.264, que pode estender-se a cobrir a televisão 4K, mas não foi projetado com a funcionalidade, o H.265 foi construído para corresponder às capacidades de telas de futuras tecnologias e inclui suporte para cores de 10 *bits* e altas taxas de quadros.

O tamanho dos arquivos gerados para transmissão de vídeo na internet é o maior problema nas transmissões em tempo real. Como exemplo, se não for feito qualquer tipo de compressão, 5 minutos de sinal de televisão com tamanho 720x576 a 25 imagens por segundo no formato 4:2:0, que necessita de 124 Mbps por segundo, ocupam a quantidade de 4,54 Gbytes (ALMEIDA, 2004).

Sendo assim, a transmissão desses arquivos pela internet, nas atuais condições da rede, tornaria a aplicação em tempo real inviável. Por isso, é grande a relevância do desenvolvimento de *codecs* que reduzam o tamanho desses arquivos sem perder informação e qualidade.

A principal motivação para a utilização do particionamento baseado em blocos em compressão de vídeo ou imagens é a possibilidade de codificar cada bloco com uma configuração específica escolhida dentro de um conjunto de parâmetros pré-definidos, considerando que, em geral, um modelo único não consegue mapear eficientemente as propriedades de uma imagem completa.

O codificador/decodificador HEVC utiliza a abordagem de particionamento Quadtree com estruturas que armazenam a informação de subdivisão de blocos de diferentes tipos. Essas estruturas servem tanto para a subdivisão dos blocos como para a parametrização para as etapas de transformação e predição do codificador. (BELTRÃO, 2011).

3.5 EFICIÊNCIA DOS CODECS

Os *codecs* tentam melhorar a eficiência explorando ao máximo as limitações humanas descritas acima, enviando a maior quantidade de informação numa transmissão de vídeo mais eficaz.

Na eliminação de redundância em codificação de vídeo são utilizadas diferentes técnicas tais como:

- Compensação de Movimento
- Transformada Discreta de Cos - seno (DCT)
- Quantificação e Codificação por VLC
- Organização do *bitstream*

3.5.1 Compensação de Movimento

Esta técnica universalmente utilizada em sistemas de vídeo. E num quadro de vídeo a quantidade de *pixels* se mantém quase inalterada entre uma imagem e próxima, produzindo uma redundância temporal entre as imagens consecutivas de

um trecho de vídeo. A partir desse fundamento que surgiu a idéia de reduzir tal redundância ao máximo e enviar uma nova imagem com o mínimo de diferença da imagem anterior.

Assim, esta técnica idealiza a criação de um *codec* cada vez mais eficiente que permita construir a próxima imagem tendo como base a imagem anterior, guardada num *buffer*.

3.5.2 Transformada Discreta de Cos-seno (DCT)

Uma técnica eficiente utilizando a tratamento matemático da informação de uma imagem do vídeo a codificar, basicamente tenta reduzir a informação eliminando ao máximo a redundância entre espaços da imagem aplicando esta técnica a cada bloco de 8×8 *pixels*. A transformada é a chave para o processo de compressão: ela toma um conjunto de pontos no domínio espacial e os transforma em uma representação equivalente no domínio da frequência.

Outro aspecto relevante é que os novos dados devem exigir menos espaço ao mesmo tempo em que o algoritmo para obtê-los seja o mais eficiente possível. A transformada utilizada no processo de compressão é a Transformada Discreta do Cosseno (TDC), que converte um bloco de pixels em uma matriz de coeficientes, descorrelacionando a informação da imagem. Os coeficientes (transformados) iniciais do bloco contêm as informações mais importantes da imagem, assim, deve-se garantir o armazenamento de certo número de coeficientes com baixos índices de posição. Por outro lado, como a transformada fornece uma série de cossenos (que no limite é convergente), os coeficientes diminuem em amplitude conforme os índices crescem. Pode-se, portanto, quantizar os coeficientes por zona, diminuindo-

os e mesmo eliminando os menos significantes, obtendo o principal ganho na taxa de compressão (ALMEIDA, 2014).

3.5.3 Quantificação e Codificação por VLC

Esta técnica combinada com a técnica Transformada Discreta do Cos-seno (DCT), na qual a informação digital que apenas pode ter um determinado número de valores. A escolha do número de valor de quantificação será tanto maior quanto maior for a qualidade da imagem reproduzida. A construção do *bitstream* final é feita através da sequência em escada Zig-zag desde o coeficiente seguinte à componente DC (coeficiente 1,1) até ao inferior direito: O coeficiente no topo superior esquerdo corresponde ao componente DC, resumindo, a energia média ou luminância da imagem e é tratado por separado.

A codificação destes pares é feita com Códigos de Comprimento Variável (VLC).

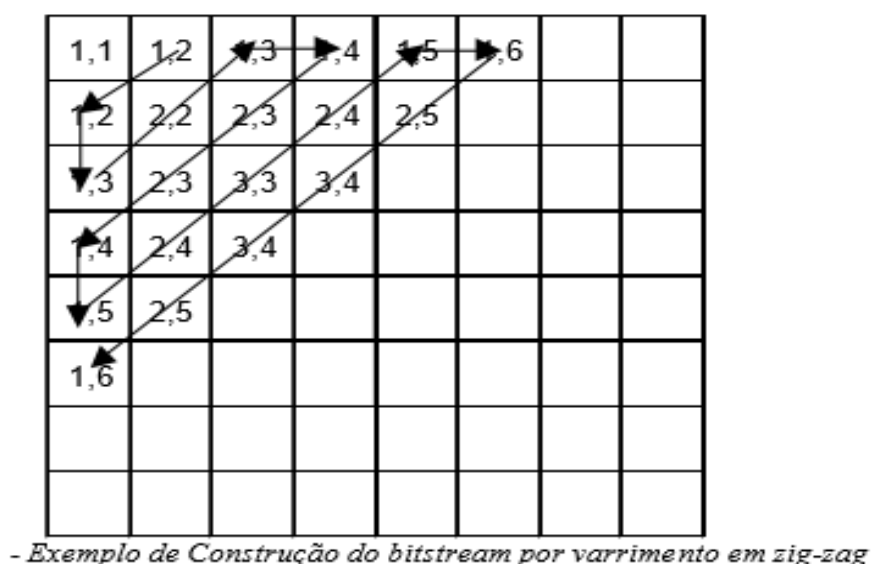


Fig. 3.5 Construção do *bitstream* por variante em zig-zag

Fonte: (Almeida, 2004)

Esta técnica gera sequências aleatórias com taxas de *bits* variáveis, comumente acoplamos um *buffer* para manter uma regularidade ao enviar a informação em qualquer canal de transmissão.

3.5.4 Organização do Bitstream

A soma resultante dos *bits* elementares de áudio e de vídeo de uma transmissão terá saída no codificador, a ordem crescente de complexidade cria uma estrutura hierárquica em: Blocos, Macro-blocos, Fatias, Quadros, Grupo de Quadros, Sequências.

- **Bloco:** O Bloco é um conjunto de 8x8 pixels
- **Macro-blocos:** Produto compensação de movimento e quantificação de um bloco de 16x16 pixels (4 blocos).
- **Fatias:** Agregação vários macro-blocos
- **Quadros:** Representa uma imagem.
- **Grupo de Quadros:** Conjunto de Quadros
- **Sequências:** Conjuntos de vários grupos de quadros (trecho de vídeo), neste caso o decodificador tem a missão de interpretar as informações que circulam nesses fluxos, sendo fornecidos dados de localização temporal, mas que na prática, podem nem sequer utilizar todos os tipos de quadros ou de fatias.

4. O PADRÃO HEVC

São descritos nas próximas seções os principais pontos da arquitetura do padrão *HEVC* com destaque para as novas funcionalidades, sua estrutura e as operações típicas para a produção do fluxo de bits válido.

4.1 Conceito de *HEVC* e suas Características

O *HEVC* (*High Efficiency Video Coding*) trabalho de desenvolvimento mútuo entre as organizações encarregadas da padronização ITUT VCEG (*Video Coding Experts Group*) e a ISO/IEC MPEG (*Moving Picture Experts Group* como JCT-VC (*Joint Collaborative Team on Video Coding*).

O padrão *HEVC* é projetado para alcançar várias metas, incluindo a alta eficiência de codificação, facilidade de integração de sistemas de transporte e resiliência em perda de informações, assim como a adaptabilidade para arquiteturas de hardware com processamento paralelo (BELTRÃO, 2011).

Este novo padrão, o *High Efficiency Video Coding* (HEVC), proporciona um aumento de quase 50% na eficiência de compressão sobre o padrão H.264/AVC, mantendo o mesmo nível de qualidade visual percebida pelo usuário final. Atendendo às necessidades que comumente foram atendidas pelo H.264/MPEG-4 AVC e, principalmente, obtendo resultados importantes em dois requisitos pontuais:

- Codificação de vídeos com alta resolução
- Uso de arquiteturas de processamento paralelo.

4.2.1 Unidade de Codificação (UC)

A Unidade de Codificação define-se como a unidade básica de processamento, o funcionamento assemelhasse muito ao do macro bloco no H.264/AVC, e contém uma ou várias UPs (Unidade de Predição) e UTs (Unidade de Transformada). Ela passa a ser subdividida em quatro blocos idênticos, até a última camada (profundidade máxima). Uma grande vantagem é o fato da UC escolher de forma aleatória os tamanhos da UC, permitindo ao codec otimizar para diferentes tipos de conteúdo, aplicações e dispositivos.

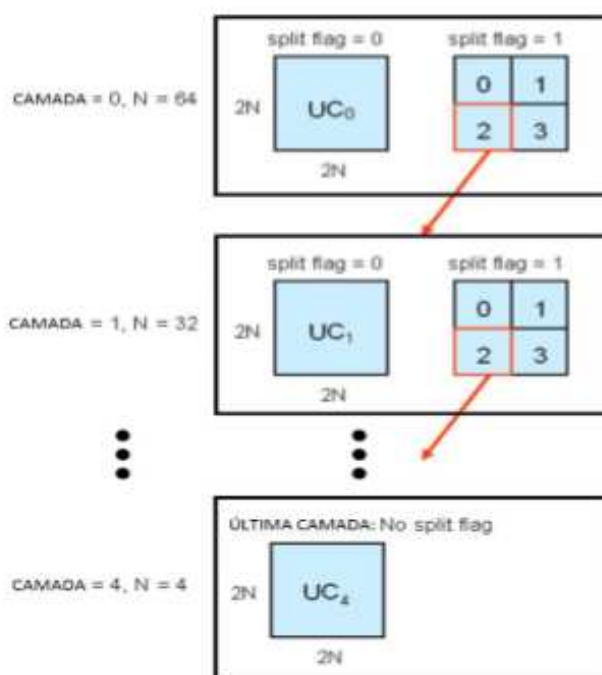


Fig. 4.7: Divisão recursiva da UC.

Fonte: (Beltrão, 2011)

Quando o processo de divisão é feito, métodos de predição são especificados na ultima camada para cada UC se dividir mais.

4.2.2 Unidade de Predição (UP)

A UP é a unidade básica de predição e carrega todas as informações relacionadas a este processo. O formato de particionamento das UPs é definido de acordo com o tipo de predição.

4.2.3 Unidade de Transformada (UT)

A Unidade de Transformada (UT) é a unidade básica do processo de transformada e quantização.

O tamanho varia de 4x4 até 32x32 amostras de luminância (não excedendo o tamanho da UC que a contém), e subdividi-se de acordo com a configuração de codificação. São definidos até 34 modos de predição intra-frame: 33 direcionais e um modo DC, a depender do tamanho da PU. Para 4x4, 8x8, 16x16, 32x32 e 64x64, existem 17, 34, 34, 34 e 5 modos disponíveis respectivamente.

Por outro lado, a compensação de movimento é uma técnica muito relevante para um ótimo desempenho do codec. O *HEVC* similar ao H.264/AVC faz uso de imagens já decodificadas e as toma como referência. Pode ser usar mais de uma imagem como referência para a codificação, assim como bi-predição. Os coeficientes dos filtros são otimizados para minimizar o número de operações aritméticas (adições e deslocamentos). Para as amostras de croma é utilizada interpolação bi-linear com precisão de 1/8-pixel.

4.2.4 Unidade Árvore de Codificação (CTU)

O CTU tem um tamanho selecionado em tempo real pelo codificador e pode ser maior ou menor que um macro-bloco tradicional. Um CTU consiste em um bloco de codificação em árvore (CTB – *Coding Tree Block* – de informações de luminância

e os CTBs de crominância correspondentes e os elementos sintáticos. O tamanho $L \times L$ de um CTB de luminância pode ter L igual a 16, 32 ou 64 amostras, com os tamanhos maiores tipicamente proporcionando melhor compressão. Dessa forma o HEVC permite um particionamento dos CTBs em blocos menores usando uma estrutura de árvore baseada no algoritmo Quadtree. Cada CTB é a raiz de uma árvore de codificação, que é utilizada para dividir a CTB em blocos de codificação (CB). O seu tamanho pode ser escolhido de forma adaptativa usando uma segmentação baseada em uma árvore Quadtree com as folhas da Quadtree representando os CBs. Cada CB é uma raiz de uma árvore de previsão e uma árvore de transformação. Cada árvore de previsão tem apenas um nível e descreve um bloco de codificação (CB – CodingBlock) ainda pode ser dividido nos chamados blocos de predição e blocos de transformação (DE OLIVEIRA, 2014).

4.3 Quadros

Um quadro é composto por *Slices*, que por sua vez são fatias do quadro, tais fatias independem uma da outra, sendo que um defeito numa, não terá afetado as outras, permitindo dessa forma serem codificadas em paralelo, o fato de dividir o quadro em *Slices* aumenta o número de bits mínimos necessários para codificá-lo (KIM, 2012 apud TEIXEIRA, 2014, p.25).

Sendo assim, o codificador pode optar por não utilizar todos os *Slices*, podendo utilizar somente de um para o quadro inteiro, a fim de agregar estas qualidades ao vídeo.

5. COMPARAÇÃO DO H.265 COM O SEU ANTECESSOR

Analizados os elementos da estrutura do codec HEVC, temos os elementos necessários para analisar e comparar os *codecs* H.265 com o *codec* mais utilizado no momento, o codec H.264, citamos as condições de restrição, elaboramos uma tabela comparativa de ambos os *codecs*.

5.1 Comparação dos *Codecs* de Padrões Atuais com os Novos

O *HEVC* (H.265) foi desenvolvido essencialmente para atender a todas as aplicações do padrão H.264/AVC e para focar em duas áreas chaves: Altas resoluções de vídeo e o melhor aproveitamento de recursos de arquiteturas de processamento paralelo (DE OLIVEIRA, 2014).

O H.264/AVC embora muito popular, mas devido às demandas de novas tecnologias como 4K e ao crescimento exponencial dos serviços de *streaming*, que vem se popularizando cada vez mais e que inevitavelmente se tornaram os responsáveis do tráfego elevado nas redes. Apresenta diferentes pontos desfavoráveis em relação ao seu sucessor o formato *HEVC* ou H.265. O HEVC é um formato de compressão de vídeo que apresenta até 40% de bits a menos em relação a codificação de um vídeo de qualidade no formato H.264 /AVC. No formato HEVC as CTUs podem ter tamanho de 8x8, 16x16, 32x32 ou 64x64, uma CTU análoga no formato H.264/AVC eles são restritos até 16x16.

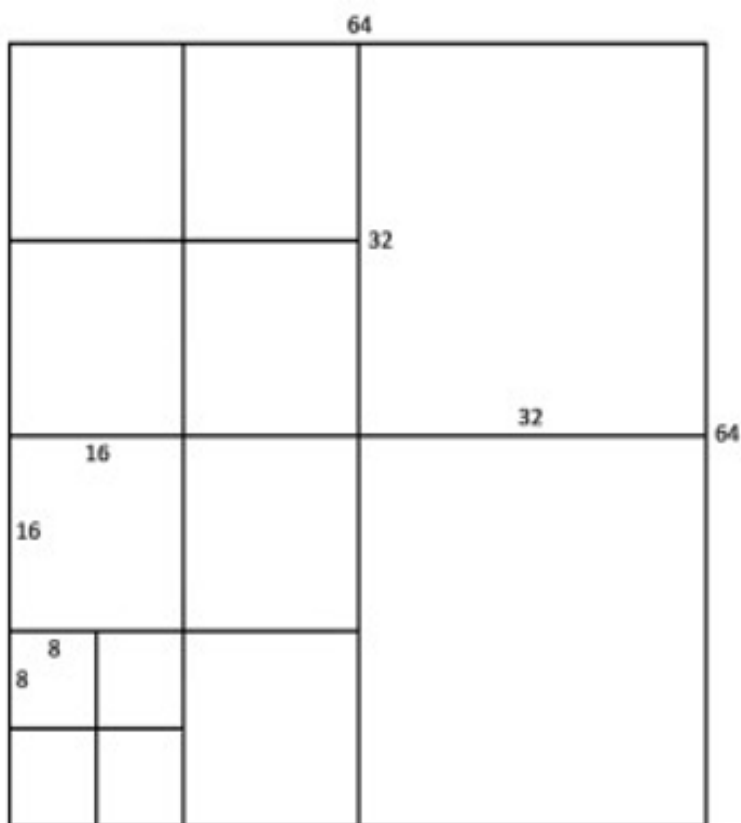


Fig. 5.8: Ilustração da divisão de uma CTU em CUs

Fonte: (Kim, 2012 apud Teixeira, 2014)

5.2 Condições de Restrição

O comitê do JCT-VC especifica seis configurações de referência para o *HEVC*. São duas condições de restrição – que representam dois cenários de aplicação:

5.2.1 *Random Access (RA)*

Para aplicações de broadcast, por exemplo, com pontos de acesso aleatório durante a sequência (1s).

5.2.2 Low Delay (LD)

Para aplicações em tempo real, sem reordenamento de imagens. A sua vez usam três modos de codificação:

- Alta eficiência (HE)
- Baixa complexidade (LC)
- Modo Intra (I)

Fornecendo um total de seis configurações: (Intra HE, Intra LC, RAHE, RA LC, LD HE e LD LC).

As figuras mostram algumas das curvas de desempenho para ambos os codecs. O Formato HEVC supera o formato H.264/AVC em todas as configurações, inclusive para as de baixa complexidade (LC), atingindo, em alguns casos, mais de 2 dB de ganho.

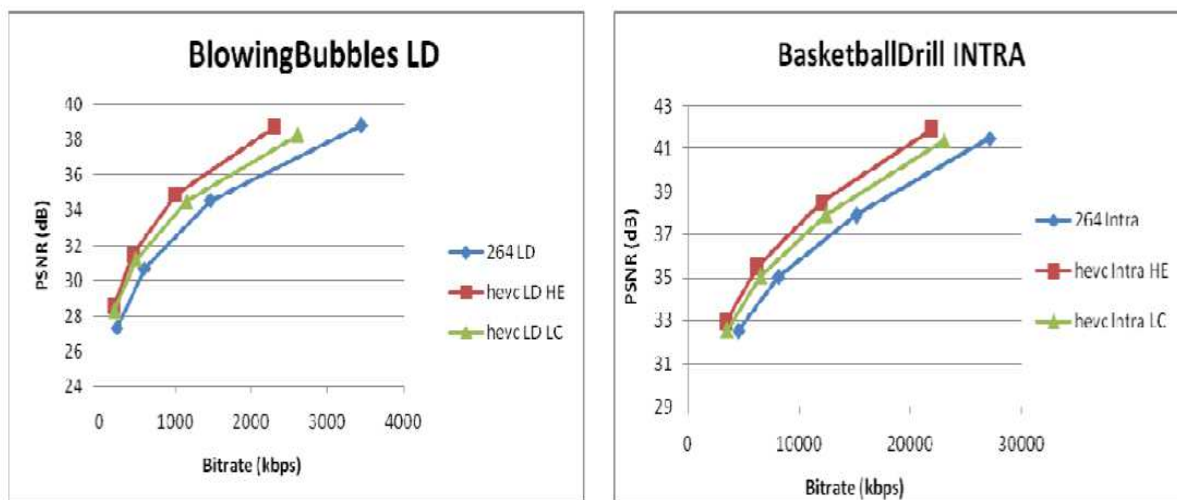


Fig. 5.9

PSNR x BitRate – BlowingBubbles – 416x240

PSNR x BitRate – BasketballDrill – 832x480

Low Delay

INTRA.

Fonte: (Beltrão, 2011)

Definitivamente o formato *HEVC* apresentou um desempenho muito superior ao formato H.264/AVC, em todas as configurações, ao custo de uma maior complexidade de codificação (DE OLIVEIRA, 2014).

Tabela: 5.2: Comparação de desempenho entre *Codec* H.264 e H.265.

Fonte: mediaentertainmentinfo.com/2013

Categoria	H.264	H265
Nomes	MPEG 4 parte 10 AVC (introduzida em 2004)	MPEG-H, HEVC, Parte 2 (Aprovado janeiro de 2013)
Adoção na Indústria	Codec de vídeo dominante e Aceito parágrafo terrestre, cabo, satélite e IPTV Transmissão. (ATSC / DVB / ISDB) Amplamente USADO em Blu-Ray, Sistemas de Segurança, videoconferência, vídeo Móvel, os jogadores de Mídia, vídeo chat, etc.	Demonstração de implementação em NAB, IBC e outros eventos, A Partir de 2012 de Empresas, Por Exemplo ATEME, Broadcom, Thomson, harmônica (Cisco), Ericsson, Qual commetc
Melhorias	- 40-50% de Redução de Taxa de pedaços em comparação com MPEG2 - liderou O Crescimento de Entrega de Conteúdo HD de Transmissão e on-line	- 40-50% da Redução da Taxa de bits deem Qualidade visuais em comparação com H.264 - Potencial parágrafo perceber UHD, 2K, 4K parágrafo Difusão e Online (OTT)
Progresso	Sucessor de MPEG-2	Sucessor do MPEG 4 AVC, H.264

Categoria	H.264	H265
Modelo de Compressão	Híbrido modelo espacial e temporal Previsão - Flexível Partição do Bloco Macro (MB), sub MB para uma Estimativa de Movimento - Previsão Intra (extrapolar Já decodificado pixels Vizinhos Quadros de predição) - Blocos de macro Estrutura com Tamanho Máximo de 16x16 - Codificação de Entropia e CABAC e CAVLC	Reforça o modelo espacial e temporal híbrido - particionamento Flexível, introduz Codificação Árvore Unitária (Codificação, Previsão e Transformação Unidades -Cu, PU, TU) - 35 Modos de Direção predição intra - Superior Arquitetura de Processamento Paralelo, Melhorias em multimídia de codificação de Extensão - CTU Estrutura de Suporte de Bloco maior (64x64), Com Mais Estruturas de Partição sub Variável - Codificação de entropia, portanto CABA.
Especificações	Suporta Até 4K (4096 × 2304) Suporta Até 59.94 fps 21 Perfis; 17 níveis	Até 8K UHD TV (8192 × 4320) Suporta ATE 300 fps 3 Perfis aprovada, Projeto Adicional de 5; 13 níveis

6. PERSPECTIVAS DE EVOLUÇÃO DO H.265

Ressaltamos as características principais do novo padrão H.265, como todas as expectativas que o mesmo promete nas novas tecnologias na qualidade de compactação de vídeo, citamos também a possibilidade de uniformizar o uso de um *codec* único no futuro.

6.1 O Potencial do H.265

A evolução da transmissão e reprodução de vídeos em tempo real está associada ao aumento de qualidade de transmissão e em dois aspectos fundamentais, o vídeo em 3D e a interatividade com o usuário, ambos visando proporcionar uma sensação de realidade (OLIVEIRA, 2014).

O Novo Padrão H.265 de codificação de vídeo, Traz promessa de economiza muito na Largura de banda, aproximadamente de 40-45% em relação ao conteúdo anteriormente codificado pelo H.264 com melhor qualidade e menores taxa de bits.

Além disso, *HEVC* tem potencial de impactar significativamente em outras áreas, incluindo aceleração da demanda e popularização de telas 4K, que tem sido amplamente limitado devido à ausência de conteúdo UHD.

Dessa forma, *HEVC* pode compensar o grande desafio de manter taxas de bits reduzidas na compreensão dos arquivos multimídia permitindo assim, as empresas de Rádio difusão e fornecedores OTT (*Over The Top*) ou serviços de Streaming na própria rede mundial de computadores agregar mais canais interativos, aumentando muito a experiência de qualidade de vídeo sem sobrecarregar muito a largura de banda como era anteriormente em relação ao seu antecessor H.264.

Atualmente, a maior percentagem de tráfego na internet provém do vídeo, dados retirados através do relatório da Cisco. E o responsável por grande parte desta fatia é exatamente o *streaming* de vídeo em tempo real (CISCO 2014).

Sendo assim, num futuro próximo pensa-se que o recurso a serviços *Pay Tv* deixe de ser procurado pelas vantagens demonstradas nos serviços OTT. O elevado uso de sistemas móveis também obriga a que num futuro próximo se uniformize os *codecs* de *stream* implementados nos diferentes sistemas. Quando se fala na evolução do *live streaming* rapidamente se pensa no aumento de qualidade de transmissão e em dois aspectos fundamentais, o vídeo a 3D e a interatividade com o utilizador, ambos visando proporcionar uma sensação de realidade.

Por isso, outra ideia de evolução passa pela interatividade com o utilizador, dando-lhe a hipótese de escolher qual o plano da câmara que pretende assistir, havendo a hipótese de poder escolher mais que um plano de visualização ao mesmo tempo, serviço *multi-view*. Os planos distintos são captados por diferentes câmaras, e todos codificados em computadores separados. O áudio é captado em separado e a sua codificação é efetuada num dos computadores utilizados para a codificação de uma das câmaras.

Numa transmissão em tempo real, tende ser garantido que mesmo quando o servidor está trocando de plano, o áudio continua a ter uma transmissão contínua.

Logo, a seleção dos planos que cada utilizador pretende visualizar é efetuada no *buffer* do servidor. Assim, o *stream* recebido pelo utilizador pode conter um ou mais planos, e no terminal é selecionado o plano a visualizar.

Todas essas futuras novidades tecnológicas relativas ao ambiente multimídia, como teleconferência em tempo real com maior qualidade e sem sobrecarregar

muito o tráfego de rede devido à grande capacidade de compressão dos arquivos serão desenvolvidos juntamente novas tecnologias para acompanhar tal novidade.

7. CONCLUSÃO

A utilização da transmissão de vídeo em tempo real na Internet está recebendo uma demanda muito grande, os arquivos a serem transferidos pela rede se tornaram maiores em tamanho e quantidades, facilitado pelo aumento da disponibilidade de *links* de internet com bandas maiores (como exemplo maiores ofertas dos *links xDSL*). Isso vem se tornando um agravante para aplicações em tempo real, que precisam evitar ao máximo o atraso na chegada dos pacotes transferidos pela rede.

Os *codecs* possuem um papel importante para a redução dos pacotes que serão transferidos na rede; este trabalho mostrou as dificuldades encontradas na rede para se manter uma bem sucedida transmissão em tempo real. Foram descritas características do padrão atual de *codec* de vídeo, o novo padrão e foi feita uma comparação do novo com o seu antecessor, para se ter uma real noção da qualidade conseguida.

As perspectivas para o futuro em relação ao desenvolvimento de melhorias nos *codecs* estão relacionadas com a vontade dos usuários de se obter mais realismo nas imagens e vídeos e com isso, em paralelo o avanço no desenvolvimento das extensões dos padrões para vídeos 3D.

Com o evoluir das tecnologias, melhoramentos ao nível de *hardware* e *software*, permitem-se um aumento de velocidade de ligação tornando possível haver *upload* e *download* e tráfego de maior qualidade num menor período de tempo.

O *live streaming* também tornou-se cada vez mais popular dando novas oportunidades de negócio. Ao nível da arquitetura, após uma abordagem e análise do modelo P2P, torna-se necessário adaptar a rede para uma rede de distribuição

com menos lacunas. Prevê-se que o protocolo *MPEG-DASH* venha a ser adotado globalmente, onde proporcionará uma distribuição de sistemas adaptativos de forma eficiente e eficaz. Quanto ao *codec* de vídeo a utilizar, embora o H.264 seja satisfatório, espera-se que o seu sucessor H.265 venha a apresentar especificações que o melhorem e por consequência seja este que será adotado globalmente.

8 TRABALHOS FUTUROS

O padrão H. 265 também conta com técnicas para previsão de movimento, detecção de redundâncias e outros aprimoramentos que permitem maior agilidade na comunicação em vídeo. A compressão inteligente das imagens favorecerá também outra tendência: a dos serviços de videoconferência pela nuvem que, assim como outras aplicações de cloud computing, eliminam o investimento em hardware e licenças.

Apesar da aparente evolução do formato H 265, que é de propriedade da MPEG LA, fabricantes de hardware e software, como a própria Cisco (que é parceira da Teltec Solutions) estão enfrentando problemas relacionados a licenças. Isso porque os termos de licenciamento impedem o uso de H.265 em qualquer tipo de código aberto ou software aplicativo distribuído gratuitamente, e em produtos freemium , como o WebEx ou Cisco Spark. Por isso, apesar de as soluções de telepresença da Cisco serem compatíveis e usufruírem dos benefícios do H.265, o codec de vídeo não pode ser considerado um formato universal como foi seu antecessor.

Nos trabalhos futuros serão explorados estudos relacionados aos novos codecs de vídeo open source e livre de pagamento de royalties que estão sendo criados para terem uma eficiência semelhante conseguida pelo H 265.

REFERÊNCIAS

ALMEIDA, Rui Jorge dos Santos. **"Transmissão de vídeo em tempo real na internet."** 2004. Disponível em: <https://ria.ua.pt/handle/10773/1862>. Acessado em 14 dezembro 2015.

BRANDI, F. **Super-resolução utilizando quadros chave em sequências de vídeo de resolução mista.** Dissertação de Mestrado em Engenharia Elétrica com ênfase em Telecomunicações, Publicação PPGENE.DM 362A/09, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 90p., 2009.

BELTRÃO, Gabriel; ARTHUR, Rangel; IANO, Yuzo. **Introdução ao HEVC–High EfficiencyVideoCoding,** 2011. Disponível em: http://www.sps.fee.unicamp.br/sps2011/proceedings_sps2011/Gabriel_Coding_SPS2011.pdf. Acessado em 10 dezembro2015.

BRUN A., Marta E. Gonçalves Vogt, Silveira A. M., **QoS – Qualidade de Serviço em TCP/IP,** 2002. Cbpf, UDP – ASPECTOS DE SEGURANÇA, Centro Brasileiro de Pesquisa Física.

CISCO.**Cisco Visual Networking Index: Forecast and Methodology,** 2014–2019. Disponível em: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.pdf. Acessado em 08 janeiro 2016.

DIAS, A. F., **Concepção Conjunta Hardware/Software de Sistemas Embarcados de Processamento de Imagens,** CDTN/CNEN, Belo Horizonte, MG.2001.

FONSECA, T. A. da.**Redução de Complexidade na Compressão de Vídeo de Alta Resolução.** Dissertação de Mestrado em Engenharia Elétrica com ênfase em Telecomunicações, Publicação PPGENE.DM 323/08, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 78p., 2008.

GOPALAKRISHNA, S.; Hannuksela, M.M.; Gabbouj, M., **"Flexible Coding Order for 3D video extension of H.265/HEVC," *Picture Coding Symposium (PCS), 2013,*** vol., no., pp.253, 256, 8-11 Dez. 2013. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6737731&isnumber=6737661>. Acessado em 15 dezembro 2015.

HERSENT, O.; Guide, D.; Petit, J. P. **Telefonia IP: comunicação multimídia baseada em pacotes.** Prentice Hall, 2002, São Paulo, Brasil.

KUROSE, J. F. e ROSS, K. - **Redes de Computadores e a Internet - 5ª Edição,** Pearson, 2010.

MACCHIAVELLO, B. L. E. (2009). **Codificador Distribuído de Vídeo com complexidade variável a partir de codificação em resolução espacial mista.** Tese de Doutorado em Engenharia Elétrica com ênfase em Telecomunicações,

Publicação MTARH. DM - 326 A/08, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 131p.

OLIVEIRA, Jean Felipe Fonseca de; DE ALENCAR, Marcelo Sampaio. **Padrão HEVC–Novas Tecnologias para Aplicações de Elevadas Taxas de Compressão de Vídeo**, 2014.

ROESLER V., **Transmissão multimídia em redes de computadores**, UNISINOS, 2009.

SANTANA H. **A camada de transporte tem como uma das principais funções a ampliação da qualidade de serviço (Quality of Service – QoS)**, Universidade Santa Cecília – Unisanta, 2005.

SILVA, Marco Antônio da. **PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO. 2005**. Tese de Doutorado. Universidade Federal de Santa Catarina.

TANENBAUM, Andrew S.; J. Wetherall: **Redes de Computadores - 5a edição**. PEARSON EDUCATION - BR, Rio de Janeiro, 2011.

TSCHÖKE, Clodoaldo. **Criação de Streaming de Vídeo para Transmissão de Sinais de Vídeo em tempo real pela internet**, 2001. Disponível em: <http://campeche.inf.furb.br/tccs/2001-II/2001-2clodoaldotschokevf.pdf>. Acessado em 10 jan. 2016.

VASCONCELOS. Laércio. **Transmissão de vídeo em redes**, 2010. <http://www.laercio.com.br/transmissao-de-videos-em-redes/>. Acessado em 20 de janeiro de 2016.

http://www.teleco.com.br/tutoriais/tutorialqosqoe/pagina_5.asp Acessado em 15 janeiro de 2016.

<http://www.mediaentertainmentinfo.com/2015/06/5-reason-2015-is-the-year-of-ott.html/> Acessado em 24 de janeiro de 2016.

<http://www.inf.ufes.br/~pdcosta/ensino/2009-1-estruturas-de-dados/material/CodificacaoHuffman.pdf> - Acessado em 24 de janeiro de 2016.